BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 11:16:03 ; Search time 221.672 Seconds
(without alignments)
10331.847 Million cell updates/sec

Title: US-09-807-933B-8
Perfect score: 1017
Sequence: 1 atgaagttcaccgttgctat.....caggttgcgaaagaaagtaa 1017

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1017	100.0	1017	21	AAA62729	Endoglucanase nucl
2	1017	100.0	1017	24	AAL43247	Rhizopus arrhizus
3	940	92.4	1164	21	AAA62730	Endoglucanase nucl
4	940	92.4	1164	24	AAL43248	Rhizopus arrhizus
5	417.6	41.1	1017	21	AAA62726	Endoglucanase nucl
6	417.6	41.1	1017	24	AAL43244	Rhizopus arrhizus
7	404.6	39.8	1083	21	AAA62728	Endoglucanase nucl
8	404.6	39.8	1083	24	AAL43246	Rhizopus arrhizus
9	393.6	38.7	1041	21	AAA62731	Endoglucanase nucl
10	393.6	38.7	1041	24	AAL43249	Phycomyces nitens
11	367.4	36.1	1101	21	AAA62727	Endoglucanase nucl
12	367.4	36.1	1101	24	AAL43245	Rhizopus arrhizus
13	288.4	28.4	1043	21	AAA62732	Endoglucanase nucl
14	288.4	28.4	1043	24	AAL43250	Rhizopus arrhizus
15	221.4	21.8	984	19	AAV16105	Fusarium oxysporum
16	221.4	21.8	1473	12	AAQ14857	Fusarium oxysporum
17	221.4	21.8	1473	13	AAQ26407	Fusarium oxysporum
18	221.4	21.8	1473	13	AAQ26382	Endoglucanase #2.
19	221.4	21.8	1473	13	AAQ25933	Cellulase containe
20	221.4	21.8	1473	13	AAQ29935	Endoglucanase gene
21	221.4	21.8	1473	14	AAQ49942	Endoglucanase enzy
22	221.4	21.8	1473	16	AAZ60179	F. oxysporum endog
23	221.4	21.8	1473	19	AAV16103	Fusarium oxysporum
24	219.8	21.6	1473	14	AAQ41733	Dye transfer inhib
25	207.4	20.4	1423	17	AAT39049	cDNA encoding cell
26	204	20.1	915	19	AAV15075	Hybrid DNA compris
27	202.8	19.9	922	19	AAV15073	Hybrid DNA compris
28	200.6	19.7	928	19	AAV15074	Hybrid DNA compris
29	188	18.5	925	19	AAV15076	Hybrid DNA compris
30	187.6	18.4	672	24	AAL43263	Humicola insolens
31	187.6	18.4	672	24	AAI69425	Humicola insolens
32	186.8	18.4	922	19	AAV15072	Hybrid DNA compris
33	186	18.3	1154	17	AAT39048	cDNA encoding cell
34	184	18.1	1174	17	AAT39050	cDNA encoding cell
35	184	18.1	1174	19	AAV39096	Monocomponent endo
36	181	17.8	807	19	AAV16104	Humicola insolens
37	179.4	17.6	1058	13	AAQ26405	Humicola insolens
38	179.4	17.6	1060	12	AAQ14856	Humicola insolens
39	179.4	17.6	1060	13	AAQ26380	Endoglucanase #1.
40	179.4	17.6	1060	13	AAQ25932	Cellulase containe
41	179.4	17.6	1060	13	AAQ29934	Endoglucanase gene
42	179.4	17.6	1060	13	AAQ30067	Sequence encoding
43	179.4	17.6	1060	14	AAQ41732	Dye transfer inhib
44	179.4	17.6	1060	14	AAQ49941	Endoglucanase enzy
45	179.4	17.6	1060	16	AAZ60178	H. insolens endogl

ALIGNMENTS

RESULT 1
AAA62729
ID AAA62729 standard; DNA; 1017 BP.
XX
AC AAA62729;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 4.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Mucor circinelloides.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09824.
XX
PT Endoglucanases of fungal origin with high activity under alkaline

PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 118-119; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1017 BP; 233 A; 255 C; 236 G; 293 T; 0 other;

Query Match 100.0%; Score 1017; DB 21; Length 1017;
Best Local Similarity 100.0%; Pred. No. 1.1e-289;
Matches 1017; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAAGTTCACCGTTGCTATTACTTCAATCGCTGTTGCACTCGCTCTCAGCTCTTCTGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAGTTCACCGTTGCTATTACTTCAATCGCTGTTGCACTCGCTCTCAGCTCTTCTGCT 60

Qy         61 GAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGACCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGACCT 120

Qy        121 ACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCTCAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCTCAA 180

Qy        181 TGTCTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTCTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCT 240

Qy        241 ACCAAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACC 300

Qy        301 AAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACC 360

Qy        361 TCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACT 420

Qy        421 CGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGT 480

Qy        481 CCTGTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTGTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGT 540

Qy        541 TGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAG 600

Qy        601 CTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGT 660

Qy        661 GGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAA 720

Qy        721 GTTACCAACACCGGTGGCGATTTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTACCAACACCGGTGGCGATTTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGT 780

Qy        781 GGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGA 840

Qy        841 GCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAA 900

Qy        901 GCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTC 960

Qy        961 AAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1017
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1017
```

RESULT 2
AAL43247
ID AAL43247 standard; DNA; 1017 BP.
XX
AC AAL43247;
XX
DT 22-AUG-2002 (first entry)
XX
DE Rhizopus arrhizus endoglucanase-related coding sequence 4.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Mucor circinelloides.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15055.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Disclosure; Page 70-73; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1017 BP; 233 A; 255 C; 236 G; 293 T; 0 other;

Query Match 100.0%; Score 1017; DB 24; Length 1017;
Best Local Similarity 100.0%; Pred. No. 1.1e-289;
Matches 1017; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAAGTTCACCGTTGCTATTACTTCAATCGCTGTTGCACTCGCTCTCAGCTCTTCTGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAGTTCACCGTTGCTATTACTTCAATCGCTGTTGCACTCGCTCTCAGCTCTTCTGCT 60
```

```
Qy          61 GAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGACCT 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGACCT 120

Qy         121 ACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCTCAA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCTCAA 180

Qy         181 TGTCTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCT 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TGTCTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCT 240

Qy         241 ACCAAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACC 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ACCAAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACC 300

Qy         301 AAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACC 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACC 360

Qy         361 TCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACT 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACT 420

Qy         421 CGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGT 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGT 480

Qy         481 CCTGTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGT 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CCTGTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGT 540

Qy         541 TGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAG 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAG 600

Qy         601 CTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACCAAGCTGGATGGTGTTGT 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACCAAGCTGGATGGTGTTGT 660

Qy         661 GGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAA 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAA 720

Qy         721 GTTACCAACACCGGTGGCGATTTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGT 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 GTTACCAACACCGGTGGCGATTTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGT 780

Qy         781 GGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGA 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGA 840

Qy         841 GCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAA 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 GCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAA 900

Qy         901 GCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTC 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 GCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTC 960

Qy         961 AAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1017
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 AAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1017
```

RESULT 3
AAA62730
ID AAA62730 standard; DNA; 1164 BP.
XX
AC AAA62730;
XX
DT 25-SEP-2000 (first entry)

XX
DE Endoglucanase nucleotide sequence 5.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Mucor circinelloides.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09825.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 122-124; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1164 BP; 272 A; 289 C; 266 G; 337 T; 0 other;

Query Match 92.4%; Score 940; DB 21; Length 1164;
Best Local Similarity 98.4%; Pred. No. 6.4e-267;
Matches 949; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy          54 TTCTGCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAG 113
               ||| |  | | | || || || || |||||| ||||||| ||||||||||||||||||
Db         201 TTCCTCCTCATCATCATCATGTAGTTCCGTCTATAGTCAATGCGGTGGCATTGGATGGAG 260

Qy         114 TGGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTA 173
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         261 TGGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTA 320

Qy         174 CTCTCAATGTCTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAA 233
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         321 CTCTCAATGTCTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAA 380

Qy         234 GACATCTACCAAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGAC 293
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         381 GACATCTACCAAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGAC 440

Qy         294 AGTAACCAAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTAC 353
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         441 AGTAACCAAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTAC 500

Qy         354 TTCCACCTCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTC 413
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         501 TTCCACCTCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTC 560

Qy         414 CACAACTCGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGT 473
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        561 CACAACTCGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGT 620

Qy        474 CACTGGTCCTGTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCA 533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 CACTGGTCCTGTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCA 680

Qy        534 AAGTGGTTGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAA 593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 AAGTGGTTGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAA 740

Qy        594 TGATGAGCTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATG 653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 TGATGAGCTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATG 800

Qy        654 GTGTTGTGGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGT 713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 GTGTTGTGGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGT 860

Qy        714 TGTTCAAGTTACCAACACCGGTGGCGATTTAGGCTCTAACCACTTTGATTTGCAAATGCC 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 TGTTCAAGTTACCAACACCGGTGGCGATTTAGGCTCTAACCACTTTGATTTGCAAATGCC 920

Qy        774 CGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 CGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGG 980

Qy        834 CTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGC 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 CTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGC 1040

Qy        894 TCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCAT 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1041 TCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCAT 1100

Qy        954 GACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAA 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101 GACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAA 1160

Qy       1014 GTAA 1017
              ||||
Db       1161 GTAA 1164
```

RESULT 4
AAL43248

```
ID   AAL43248 standard; DNA; 1164 BP.
XX
AC   AAL43248;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Rhizopus arrhizus endoglucanase-related coding sequence 5.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Mucor circinelloides.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15056.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp  -
XX
PS   Disclosure; Page 75-78; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for
CC   processing fibres, de-inking waste paper and improving the freeness of
CC   paper pulp - which is particularly applicable in detergent compositions.
CC   The present DNA sequence represents an endoglucanase-related gene
CC   sequence of the invention.
XX
SQ   Sequence 1164 BP; 272 A; 289 C; 266 G; 337 T; 0 other;
```

```
  Query Match           92.4%;  Score 940;  DB 24;  Length 1164;
  Best Local Similarity   98.4%;  Pred. No. 6.4e-267;
  Matches  949;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         54 TTCTGCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAG 113
              ||| |  | | | || || || || |||||| |||||||| ||||||||||||||||||
Db        201 TTCCTCCTCATCATCATCATGTAGTTCCGTCTATAGTCAATGCGGTGGCATTGGATGGAG 260

Qy        114 TGGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTA 173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 TGGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTA 320

Qy        174 CTCTCAATGTCTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAA 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 CTCTCAATGTCTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAA 380

Qy        234 GACATCTACCAAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGAC 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GACATCTACCAAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGAC 440

Qy        294 AGTAACCAAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTAC 353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 AGTAACCAAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTAC 500

Qy        354 TTCCACCTCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTC 413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 TTCCACCTCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTC 560

Qy        414 CACAACTCGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGT 473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 CACAACTCGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGT 620

Qy        474 CACTGGTCCTGTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCA 533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 CACTGGTCCTGTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCA 680

Qy        534 AAGTGGTTGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAA 593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 AAGTGGTTGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAA 740

Qy        594 TGATGAGCTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATG 653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 TGATGAGCTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATG 800

Qy        654 GTGTTGTGGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGT 713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 GTGTTGTGGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGT 860

Qy        714 TGTTCAAGTTACCAACACCGGTGGCGATTTAGGCTCTAACCACTTTGATTTGCAAATGCC 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 TGTTCAAGTTACCAACACCGGTGGCGATTTAGGCTCTAACCACTTTGATTTGCAAATGCC 920

Qy        774 CGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 CGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGG 980
```

```
Qy        834 CTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGC 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 CTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGC 1040

Qy        894 TCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCAT 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1041 TCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCAT 1100

Qy        954 GACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAA 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101 GACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAA 1160

Qy       1014 GTAA 1017
              ||||
Db       1161 GTAA 1164
```

RESULT 5
AAA62726
ID AAA62726 standard; DNA; 1017 BP.
XX
AC AAA62726;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 1.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Rhizopus oryzae.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09821.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 104-105; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1017 BP; 240 A; 250 C; 235 G; 292 T; 0 other;

Query Match 41.1%; Score 417.6; DB 21; Length 1017;
Best Local Similarity 67.8%; Pred. No. 9.9e-113;
Matches 655; Conservative 0; Mismatches 284; Indels 27; Gaps 4;

```
Qy         64 GCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGACCTACC 123
              ||||||  || |||   | ||||||||||||||||| |    |||| ||| |||||
Db         67 GCTGCTGAATGTAGCAAATTGTATGGTCAATGTGGTGGTAAGAACTGGAATGGCCCTACT 126

Qy        124 TGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCTCAATGT 183
              |||||||||  ||| || || ||      | ||| | ||||| | ||||||||||||||||
Db        127 TGTTGTGAATCTGGATCCACCTG------TAAAGTAAGCAACGATTACTACTCTCAATGT 180

Qy        184 CTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTACC 243
              ||||||         ||||  ||||      |   ||       |||||||||
Db        181 CTTCCCTCTGGAAGCAGTGGCAATAAATCTTCTGAAAGTGCTCACAAGAAGAC------- 233

Qy        244 AAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAG 303
                 ||  || || |   |  ||| ||||| |  |||      | |||||    ||
Db        234 --TACCACTGCTGCTCACAAGAAGACTACTACCGCTGCTCATAAAAAGACTACCACTGCT 291

Qy        304 ACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCT 363
              |  |||  ||||| || ||     | |   || |  || || ||    ||  | |
Db        292 CCTGCTAAGAAGACTACAACTGTTGCCAAAGCTTCCACCCCTTCTAACTCTAGCTCTAGC 351

Qy        364 TCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGT 423
              ||    |    ||    |  ||||||| |||   ||||| |  |||  ||| ||||||
Db        352 TCCAGCGGCAAATATTCCGCTGTCTCTGGTGGTGCCTCTGGTAACGGTGTCACTACTCGT 411

Qy        424 TATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCT 483
              ||||||||||| ||||| || || || |||||||| || || ||   ||||| |  ||||
Db        412 TATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGTAAGGCCAATGTCAGTTCTCCT 471

Qy        484 GTTGACACCTGTGCCTCCAATGGTATCTCT---TTATTAGATGCCAATGCTCAAAGTGGT 540
              ||  |  |||||  |    |||||  || ||    |    ||  |||||| ||||||||
Db        472 GTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGTGACAGCAATGCCCAAAGTGGC 531

Qy        541 TGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAG 600
              |||||||||||||   ||| ||||||||| ||||||||||||||||||| || ||  |
Db        532 TGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAACCTTGGGCTGTAAACGACAAC 591

Qy        601 CTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGT 660
              || || || ||||||||||||||  | ||   |||       ||| || | ||||| |||
Db        592 CTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGTGGTGAATCTCGCTGGTGCTGT 651

Qy        661 GGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAA 720
              ||||  ||| | || |||||||||  | ||| | |||| ||||||||||||||| |||
Db        652 TCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCTGGTAAGAAGATGGTTGTCCAA 711

Qy        721 GTTACCAACACCGGTGGCGATTTAGGCTC---------TAACCACTTTGATTTGCAAATG 771
              || || ||||| ||||| ||| | |||||         |   |||||||| |||||||||
Db        712 GTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGTGCTCACTTTGACTTGCAAATG 771

Qy        772 CCCGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGAT 831
              |||||||||| ||||||||  |||||||| ||| |    |||||||| ||||||||||
Db        772 CCCGGTGGTGGTGTTGGTATTTTCAATGGTTGTTCCAGCCAATGGGGTGCTCCCAATGAC 831

Qy        832 GGCTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCT 891
              || |||||  | ||||| ||||||  |     ||||  ||||||||  |   ||||||||
Db        832 GGTTGGGGCTCAAGATACGGTGGTATTTCTTCTGCATCTGACTGCTCTAGTCTTCCTTCC 891

Qy        892 GCTCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACC 951
              || || ||||||||||||||||||||||||||||||||||||||||| |||||||||| | |
Db        892 GCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACGCTGATAACCCAAGC 951

Qy        952 ATGACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGA 1011
              ||||| | ||||||||||||||||||    ||| | ||  |     |||||||   ||||
Db        952 ATGACTTACAAGGAAGTTACCTGTCCTAAGGAAATCACCGCCAAGACAGGTTGTTCAAGA 1011

Qy       1012 AAGTAA 1017
              || |||
Db       1012 AAATAA 1017
```

RESULT 6
AAL43244
ID AAL43244 standard; DNA; 1017 BP.
XX
AC AAL43244;

```
XX
DT   22-AUG-2002  (first entry)
XX
DE   Rhizopus arrhizus endoglucanase-related coding sequence 1.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Rhizopus arrhizus.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15052.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp -
XX
PS   Example 10; Page 56-58; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for
CC   processing fibres, de-inking waste paper and improving the freeness of
CC   paper pulp - which is particularly applicable in detergent compositions.
CC   The present DNA sequence represents an endoglucanase-related gene
CC   sequence of the invention.
XX
SQ   Sequence 1017 BP; 240 A; 250 C; 235 G; 292 T; 0 other;

  Query Match             41.1%;  Score 417.6;  DB 24;  Length 1017;
  Best Local Similarity   67.8%;  Pred. No. 9.9e-113;
  Matches  655;  Conservative    0;  Mismatches  284;  Indels   27;  Gaps   4;

Qy         64 GCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGACCTACC 123
Db         67 GCTGCTGAATGTAGCAAATTGTATGGTCAATGTGGTGGTAAGAACTGGAATGGCCCTACT 126
Qy        124 TGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCTCAATGT 183
Db        127 TGTTGTGAATCTGGATCCACCTG------TAAAGTAAGCAACGATTACTACTCTCAATGT 180
Qy        184 CTTCCCGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTACC 243
Db        181 CTTCCCTCTGGAAGCAGTGGCAATAAATCTTCTGAAAGTGCTCACAAGAAGAC------- 233
Qy        244 AAGACATCTACTACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAG 303
Db        234 --TACCACTGCTGCTCACAAGAAGACTACTACCGCTGCTCATAAAAAGACTACCACTGCT 291
Qy        304 ACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCT 363
Db        292 CCTGCTAAGAAGACTACAACTGTTGCCAAAGCTTCCACCCCTTCTAACTCTAGCTCTAGC 351
Qy        364 TCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGT 423
Db        352 TCCAGCGGCAAATATTCCGCTGTCTCTGGTGGTGCCTCTGGTAACGGTGTCACTACTCGT 411
Qy        424 TATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCT 483
Db        412 TATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGTAAGGCCAATGTCAGTTCTCCT 471
Qy        484 GTTGACACCTGTGCCTCCAATGGTATCTCT---TTATTAGATGCCAATGCTCAAAGTGGT 540
Db        472 GTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGTGACAGCAATGCCCAAAGTGGC 531
Qy        541 TGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAG 600
Db        532 TGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAACCTTGGGCTGTAAACGACAAC 591
Qy        601 CTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGT 660
Db        592 CTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGTGGTGAATCTCGCTGGTGCTGT 651
Qy        661 GGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAA 720
Db        652 TCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCTGGTAAGAAGATGGTTGTCCAA 711
Qy        721 GTTACCAACACCGGTGGCGATTTAGGCTC---------TAACCACTTTGATTTGCAAATG 771
Db        712 GTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGTGCTCACTTTGACTTGCAAATG 771
Qy        772 CCCGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGAT 831
Db        772 CCCGGTGGTGGTGTTGGTATTTTCAATGGTTGTTCCAGCCAATGGGGTGCTCCCAATGAC 831
Qy        832 GGCTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCT 891
Db        832 GGTTGGGGCTCAAGATACGGTGGTATTTCTTCTGCATCTGACTGCTCTAGTCTTCCTTCC 891
Qy        892 GCTCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACC 951
Db        892 GCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACGCTGATAACCCAAGC 951
Qy        952 ATGACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGA 1011
Db        952 ATGACTTACAAGGAAGTTACCTGTCCTAAGGAAATCACCGCCAAGACAGGTTGTTCAAGA 1011
Qy       1012 AAGTAA 1017
Db       1012 AAATAA 1017

RESULT 7
AAA62728
ID   AAA62728 standard; DNA; 1083 BP.
XX
AC   AAA62728;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Endoglucanase nucleotide sequence 3.
XX
KW   Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW   animal foodstuff; ss.
XX
OS   Rhizopus oryzae.
XX
PN   WO200024879-A1.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-JP05884.
XX
PR   23-OCT-1998;   98JP-0302387.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakamura Y,  Moriya T,  Baba Y,  Yanai K,  Sumida N,  Nishimura T;
PI   Murashima K,  Nakane A,  Yaguchi T,  Koga J,  Murakami T,  Kono T;
XX
DR   WPI; 2000-365117/31.
DR   P-PSDB; AAB09823.
```

```
XX
PT   Endoglucanases of fungal origin with high activity under alkaline
PT   conditions for production of paper pulp and animal feedstuffs -
XX
PS   Claim 44; Page 113-115; 180pp; Japanese.
XX
CC   This sequence encodes an endoglucanase protein. The invention relates
CC   to an endoglucanase of fungal origin which can completely break down
CC   purified cellulose at a concentration of less than 1mg protein/litre,
CC   and produces more than 50% breakdown of cellulose at pH 8.5. The
CC   invention includes endoglucanase protein sequences (see
CC   AAB09825-B09830), endoglucanase nucleotide sequences (see
CC   AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC   identification of the endoglucanase sequences, and in the construction of
CC   vectors containing the polynucleotides. The endoglucanase enzymes are
CC   used for the production of pulp for papermaking and for the production of
CC   animal foodstuffs.
XX
SQ   Sequence 1083 BP; 260 A; 297 C; 231 G; 295 T; 0 other;

  Query Match             39.8%;  Score 404.6;  DB 21;  Length 1083;
  Best Local Similarity   65.6%;  Pred. No. 7.1e-109;
  Matches  671;  Conservative    0;  Mismatches  289;  Indels   63;  Gaps    3;

Qy         58 GCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGA 117
              ||  | ||||||   || |||   |  ||    ||||||||||| |    .|||  ||||
Db         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120

Qy        118 CCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCT 177
              |||||||| ||||||  ||||||||||||||||||| | |    | |||    | ||||||
Db        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180

Qy        178 CAATGTCTTCC---------CGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACC 228
              ||||||  ||||         |    ||  || ||||||  ||  | ||   | || |||
Db        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240

Qy        229 AAGAAGACATCTACCAAGACATCTACT-------------------------------- 255
              |   |||   | |  |||||  | |||
Db        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300

Qy        256 ------------ACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAG 303
                         |||||| | |  |   || | |     |||||||  || | | ||
Db        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360

Qy        304 ACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCT 363
              |  |  ||||  || ||||| ||||      ||||  || || | ||| | ||  | |||
Db        361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420

Qy        364 TCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGT 423
              ||||| |    |||   |   ||||||| |||   || || | ||||    || |||||
Db        421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480

Qy        424 TATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCT 483
              || ||||||||||||||  |||||||||| ||||| || || |||  ||||||    |||
Db        481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540

Qy        484 GTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGT 543
              |||| | |||||  |    ||||||   || |    |||  ||| ||||||  || |||
Db        541 GTTGGCTCCTGTAACAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGT 600

Qy        544 AACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTC 603
              |||||||   | | ||  |||||  |||| ||||||||||| |||  |  || || ||
Db        601 GTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTT 660

Qy        604 GCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGC 663
              || ||||||||| |||||  || |||  ||||    | |||||||   |||||||||| |
Db        661 GCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCC 720

Qy        664 TGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTT 723
              ||||  |||  | || ||||| |||   || |     || ||||||||||||||||||||
Db        721 TGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTA 780

Qy        724 ACCAACACCGGTGGCGATTTAGGCTCTAAC---------CACTTTGATTTGCAAATGCCC 774
              |||||||| |||   ||  | |||||||||         |||||||| ||||||||||||
Db        781 ACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCC 840

Qy        775 GGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGC 834
              ||||||| |||||||||| |||||| |||||  ||||||||| |||||||  |||||
Db        841 GGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGT 900

Qy        835 TGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCT 894
              ||||| || ||||| || |||||     |||| ||||||||||  |     |||||  |||||
Db        901 TGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCC 960

Qy        895 CTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATG 954
              |||||||||||||||| |||||||||  |||||||| ||| |||||||||| ||||||
Db        961 CTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATG 1020

Qy        955 ACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAG 1014
              |||| |||  |||||||||||||     |   | ||| |    || || ||   ||||||
Db       1021 ACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAA 1080

Qy       1015 TAA 1017
              |||
Db       1081 TAA 1083


RESULT 8
AAL43246
ID   AAL43246 standard; DNA; 1083 BP.
XX
AC   AAL43246;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Rhizopus arrhizus endoglucanase-related coding sequence 3.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Rhizopus arrhizus.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15054.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp -
XX
PS   Disclosure; Page 65-68; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for
CC   processing fibres, de-inking waste paper and improving the freeness of
CC   paper pulp - which is particularly applicable in detergent compositions.
CC   The present DNA sequence represents an endoglucanase-related gene
CC   sequence of the invention.
XX
SQ   Sequence 1083 BP; 260 A; 297 C; 231 G; 295 T; 0 other;

  Query Match             39.8%;  Score 404.6;  DB 24;  Length 1083;
```

```
Best Local Similarity   65.6%;  Pred. No. 7.1e-109;
Matches  671;  Conservative    0;  Mismatches  289;  Indels   63;  Gaps    3;

Qy         58 GCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGA 117
              ||  | ||||||   || |||    |  ||    |||||||||| |    |||  ||||
Db         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120

Qy        118 CCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCT 177
              |||||||| ||||||  |||||||||||||||||| | |    | |||    | ||||||
Db        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180

Qy        178 CAATGTCTTCC---------CGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACC 228
              ||||||  ||||         |    ||  || ||||||  ||  | ||   | || |||
Db        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240

Qy        229 AAGAAGACATCTACCAAGACATCTACT------------------------------- 255
              |   |||   | |  |||||  | |||
Db        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300

Qy        256 ------------ACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAG 303
                         |||||| | |  |   || | |     |||||||  || | | ||
Db        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360

Qy        304 ACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCT 363
              |  |  |||| || ||||| ||||       ||||  || || | ||| | ||  | |||
Db        361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420

Qy        364 TCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGT 423
              |||||  |   |||  |  |||||||  |||    || || | ||||   || |||||
Db        421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480

Qy        424 TATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCT 483
              || ||||||||||||||  |||||||||| |||||  || ||  |||  ||||||    |||
Db        481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540

Qy        484 GTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGT 543
              ||||  | |||||  |    ||||||  || |    |||  |||  ||||||  || |||
Db        541 GTTGGCTCCTGTAACAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGT 600

Qy        544 AACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTC 603
              |||||||   | | ||  |||||  |||| ||||||||||| |||  |  || || ||
Db        601 GTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTT 660

Qy        604 GCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGC 663
              || |||||||||| ||||| || ||| ||||    | ||||||  ||||||||||  |
Db        661 GCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCC 720

Qy        664 TGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTT 723
              ||||  ||| | || ||||| |||   || |     || |||||||||||||||||||
Db        721 TGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTA 780

Qy        724 ACCAACACCGGTGGCGATTTAGGCTCTAAC---------CACTTTGATTTGCAAATGCCC 774
              |||||||| |||   ||  | |||||||||         |||||||| ||||||||||||
Db        781 ACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCC 840

Qy        775 GGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGC 834
              |||||||| |||||||| |||||| |||||  ||||||||||  |||||||  |||||
Db        841 GGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGT 900

Qy        835 TGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCT 894
              |||||  || |||||  || |||||     |||| ||||||||||  |     |||||  |||||
Db        901 TGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCC 960

Qy        895 CTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATG 954
              |||||||||||||||| |||||||  |||||||| ||| |||||||||| ||||||
Db        961 CTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATG 1020

Qy        955 ACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAG 1014
              |||| |||  |||||||||||||||    | ||| |    || || ||   ||||||
Db       1021 ACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAA 1080

Qy       1015 TAA 1017
              |||
Db       1081 TAA 1083
```

RESULT 9
AAA62731
ID   AAA62731 standard; DNA; 1041 BP.
XX
AC   AAA62731;
XX
DT   25-SEP-2000 (first entry)
XX
DE   Endoglucanase nucleotide sequence 6.
XX
KW   Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW   animal foodstuff; ss.
XX
OS   Phycomyces nitens.
XX
PN   WO200024879-A1.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-JP05884.
XX
PR   23-OCT-1998;   98JP-0302387.
XX
PA   (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI   Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR   WPI; 2000-365117/31.
DR   P-PSDB; AAB09826.
XX
PT   Endoglucanases of fungal origin with high activity under alkaline
PT   conditions for production of paper pulp and animal feedstuffs -
XX
PS   Claim 44; Page 128-129; 180pp; Japanese.
XX
CC   This sequence encodes an endoglucanase protein. The invention relates
CC   to an endoglucanase of fungal origin which can completely break down
CC   purified cellulose at a concentration of less than 1mg protein/litre,
CC   and produces more than 50% breakdown of cellulose at pH 8.5. The
CC   invention includes endoglucanase protein sequences (see
CC   AAB09825-B09830), endoglucanase nucleotide sequences (see
CC   AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC   identification of the endoglucanase sequences, and in the construction of
CC   vectors containing the polynucleotides. The endoglucanase enzymes are
CC   used for the production of pulp for papermaking and for the production of
CC   animal foodstuffs.
XX
SQ   Sequence 1041 BP; 225 A; 352 C; 248 G; 216 T; 0 other;

```
  Query Match           38.7%;  Score 393.6;  DB 21;  Length 1041;
  Best Local Similarity   63.7%;  Pred. No. 1.2e-105;
  Matches  634;  Conservative    0;  Mismatches  329;  Indels   33;  Gaps    1;

Qy         55 TCTGCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGT 114
              ||  || | ||||||   ||||||   | |||||| || ||||||||||     |||| |
Db         46 TCCACTTACGCTGCTGAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACT 105

Qy        115 GGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTAC 174
              || || ||||| ||       ||||  || || || | |   |||  |||| | |  |||
Db        106 GGTCCCACCTGCTGCACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTAC 165

Qy        175 TCTCAATGTCTTCCCGGATCCCA--------------------------------CAGT 201
              ||||| ||| | |||      ||                                ||
Db        166 TCTCAGTGTATCCCCAACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCACC 225

Qy        202 AACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTACCAAGACATCTACTACCACC 261
              | ||| |||| |   | | ||   |       ||  | |||||| |  | || ||||||
Db        226 ACCAAGGCTGCCACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACC 285
```

```
Qy        262 GCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAGACAACTACCAAGACAACT 321
Db        286 ACCACCAAGGCCCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACC 345

Qy        322 ACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCTTCTTCTGCTGGTTACAAG 381
Db        346 ACCAAGACCACCACCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGC 405

Qy        382 GTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGTTATTGGGATTGTTGTAAA 441
Db        406 CCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAG 465

Qy        442 GCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCTGTTGACACCTGTGCCTCC 501
Db        466 CCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAG 525

Qy        502 AATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGTAACGGTGGTAATGGTTTC 561
Db        526 GATGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTAC 585

Qy        562 ATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTCGCTTACGGTTTCGCTGCT 621
Db        586 ATGTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCT 645

Qy        622 GCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGCTGTTATGAATTGACCTTC 681
Db        646 GCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTC 705

Qy        682 ACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTTACCAACACCGGTGGCGAT 741
Db        706 ACCAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGAT 765

Qy        742 TTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGTGGCGTTGGTATCTTCAATGGC 801
Db        766 CTCAGCACCAACCACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGC 825

Qy        802 TGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGAGCTAGATATGGTGGTGTCAGC 861
Db        826 TGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGC 885

Qy        862 TCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGATTC 921
Db        886 TCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTC 945

Qy        922 AACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTCAAGGAAGTTACCTGTCCTGCT 981
Db        946 GGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCC 1005

Qy        982 GAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1017
Db       1006 GAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041
```

RESULT 10
AAL43249
ID AAL43249 standard; DNA; 1041 BP.
XX
AC AAL43249;
XX
DT 22-AUG-2002 (first entry)
XX
DE Phycomyces nitens endoglucanase-related coding sequence.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Phycomyces nitens.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15057.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Disclosure; Page 81-83; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1041 BP; 225 A; 352 C; 248 G; 216 T; 0 other;

Query Match 38.7%; Score 393.6; DB 24; Length 1041;
Best Local Similarity 63.7%; Pred. No. 1.2e-105;
Matches 634; Conservative 0; Mismatches 329; Indels 33; Gaps 1;

```
Qy         55 TCTGCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGT 114
Db         46 TCCACTTACGCTGCTGAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACT 105

Qy        115 GGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTAC 174
Db        106 GGTCCCACCTGCTGCACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTAC 165

Qy        175 TCTCAATGTCTTCCCGGATCCCA----------------------------------CAGT 201
Db        166 TCTCAGTGTATCCCCAACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCACC 225

Qy        202 AACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTACCAAGACATCTACTACCACC 261
Db        226 ACCAAGGCTGCCACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACC 285

Qy        262 GCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAGACAACTACCAAGACAACT 321
Db        286 ACCACCAAGGCCCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACC 345

Qy        322 ACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCTTCTTCTGCTGGTTACAAG 381
Db        346 ACCAAGACCACCACCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGC 405

Qy        382 GTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGTTATTGGGATTGTTGTAAA 441
Db        406 CCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAG 465

Qy        442 GCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCTGTTGACACCTGTGCCTCC 501
Db        466 CCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAG 525

Qy        502 AATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGTAACGGTGGTAATGGTTTC 561
Db        526 GATGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTAC 585

Qy        562 ATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTCGCTTACGGTTTCGCTGCT 621
Db        586 ATGTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCT 645

Qy        622 GCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGCTGTTATGAATTGACCTTC 681
```

```
Db        646 GCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTC 705

Qy        682 ACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTTACCAACACCGGTGGCGAT 741

Db        706 ACCAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGAT 765

Qy        742 TTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGTGGCGTTGGTATCTTCAATGGC 801

Db        766 CTCAGCACCAACCACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGC 825

Qy        802 TGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGAGCTAGATATGGTGGTCTAGC 861

Db        826 TGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGC 885

Qy        862 TCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGATTC 921

Db        886 TCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTC 945

Qy        922 AACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTCAAGGAAGTTACCTGTCCTGCT 981

Db        946 GGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCC 1005

Qy        982 GAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1017

Db       1006 GAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041
```

RESULT 11
AAA62727
ID AAA62727 standard; DNA; 1101 BP.
XX
AC AAA62727;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 2.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Rhizopus oryzae.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09822.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 108-110; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1101 BP; 268 A; 258 C; 257 G; 318 T; 0 other;

Query Match 36.1%; Score 367.4; DB 21; Length 1101;
Best Local Similarity 70.9%; Pred. No. 6.8e-98;
Matches 520; Conservative 0; Mismatches 201; Indels 12; Gaps 2;

```
Qy        297 AACCAAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTC 356

Db        369 AACGACTACTACCACTGCTCCCGCTAAGGAAATTACAACTACTGCCAAAGCTTCAAACTC 428

Qy        357 CACCTCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCAC 416

Db        429 TTCTAACTCTAGCGGCAAATACTCCATTGTCTCTGGTGGTGCCTCTGGTAACGGTGTCAC 488

Qy        417 AACTCGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCAC 476

Db        489 TACTCGTTATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGTAAGGCCAATGTCAG 548

Qy        477 TGGTCCTGTTGACACCTGTGCCTCCAATGGTATCTCT---TTATTAGATGCCAATGCTCA 533

Db        549 TTCTCCTGTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGTGACAGCAATGTCCA 608

Qy        534 AAGTGGTTGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAA 593

Db        609 AAGTGGCTGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAGCCTTGGGCTGTAAA 668

Qy        594 TGATGAGCTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATG 653

Db        669 CGATAATCTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGTGGTGAATCTCGCTG 728

Qy        654 GTGTTGTGGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGT 713

Db        729 GTGCTGTTCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCTGGTAAGAAGATGGT 788

Qy        714 TGTTCAAGTTACCAACACCGGTGGCGATTTAGGCTC---------TAACCACTTTGATTT 764

Db        789 TATCCAAGTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGTGCTCACTTTGACTT 848

Qy        765 GCAAATGCCCCGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCC 824

Db        849 GCAAATGCCCGGTGGTGGTGTTGGTATTTTCAATGGTTGCTCCAAGCAATGGGGTGCTCC 908

Qy        825 CAATGATGGCTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCT 884

Db        909 CAATGACGGTTGGGGCTCGAGATACGGTGGTATTTCTTCTGCATCTGACTGCTCTAGTCT 968

Qy        885 TCCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAA 944

Db        969 TCCTTCCGCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACGCTGATAA 1028

Qy        945 CCCTACCATGACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTG 1004

Db       1029 CCCAAGCATGACTTACAAGGAAGTTACCTGTCCCAAGGAAATCACCGCCAAGACAGGTTG 1088

Qy       1005 CGAAAGAAAGTAA 1017

Db       1089 TTCAAGAAAATAA 1101
```

RESULT 12
AAL43245
ID AAL43245 standard; DNA; 1101 BP.
XX
AC AAL43245;
XX
DT 22-AUG-2002 (first entry)
XX
DE Rhizopus arrhizus endoglucanase-related coding sequence 2.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;

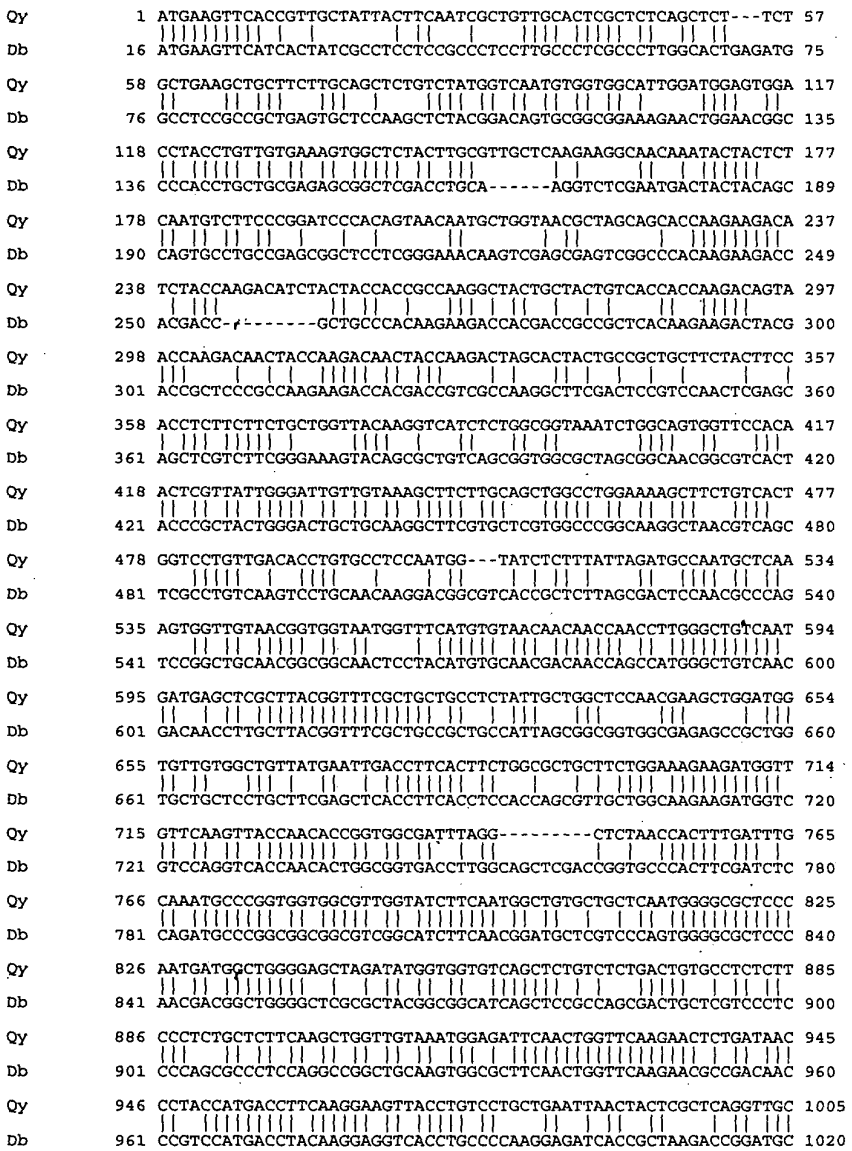
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Rhizopus arrhizus.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15053.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Disclosure; Page 60-63; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1101 BP; 268 A; 258 C; 257 G; 318 T; 0 other;

Query Match 36.1%; Score 367.4; DB 24; Length 1101;
Best Local Similarity 70.9%; Pred. No. 6.8e-98;
Matches 520; Conservative 0; Mismatches 201; Indels 12; Gaps 2;

```
Qy       297 AACCAAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTC 356
Db       369 AACGACTACTACCACTGCTCCCGCTAAGGAAATTACAACTACTGCCAAAGCTTCAAACTC 428

Qy       357 CACCTCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCAC 416
Db       429 TTCTAACTCTAGCGGCAAATACTCCATTGTCTCTGGTGGTGCCTCTGGTAACGGTGTCAC 488

Qy       417 AACTCGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCAC 476
Db       489 TACTCGTTATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGTAAGGCCAATGTCAG 548

Qy       477 TGGTCCTGTTGACACCTGTGCCTCCAATGGTATCTCT---TTATTAGATGCCAATGCTCA 533
Db       549 TTCTCCTGTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGTGACAGCAATGTCCA 608

Qy       534 AAGTGGTTGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAA 593
Db       609 AAGTGGCTGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAGCCTTGGGCTGTAAA 668

Qy       594 TGATGAGCTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATG 653
Db       669 CGATAATCTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGTGGTGAATCTCGCTG 728

Qy       654 GTGTTGTGGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGT 713
Db       729 GTGCTGTTCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCTGGTAAGAAGATGGT 788

Qy       714 TGTTCAAGTTACCAACACCGGTGGCGATTTAGGCTC---------TAACCACTTTGATTT 764
Db       789 TATCCAAGTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGTGCTCACTTTGACTT 848

Qy       765 GCAAATGCCCGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCC 824
Db       849 GCAAATGCCCGGTGGTGGTGTTGGTATTTTCAATGGTTGCTCCAAGCAATGGGGTGCTCC 908

Qy       825 CAATGATGGCTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCT 884
Db       909 CAATGACGGTTGGGGCTCGAGATACGGTGGTATTTCTTCTGCATCTGACTGCTCTAGTCT 968

Qy       885 TCCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAA 944
Db       969 TCCTTCCGCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACGCTGATAA 1028

Qy       945 CCCTACCATGACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTG 1004
Db      1029 CCCAAGCATGACTTACAAGGAAGTTACCTGTCCCAAGGAAATCACCGCCAAGACAGGTTG 1088

Qy      1005 CGAAAGAAAGTAA 1017
Db      1089 TTCAAGAAAATAA 1101
```

RESULT 13
AAA62732
ID AAA62732 standard; DNA; 1043 BP.
XX
AC AAA62732;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 7.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Rhizopus oryzae.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 132-134; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1043 BP; 212 A; 370 C; 291 G; 170 T; 0 other;

Query Match 28.4%; Score 288.4; DB 21; Length 1043;
Best Local Similarity 58.7%; Pred. No. 1.4e-74;
Matches 606; Conservative 0; Mismatches 396; Indels 30; Gaps 5;



RESULT 14
AAL43250
ID AAL43250 standard; DNA; 1043 BP.
XX
AC AAL43250;
XX
DT 22-AUG-2002 (first entry)
XX
DE Rhizopus arrhizus endoglucanase-related codon-optimised DNA sequence.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Rhizopus arrhizus.
OS Synthetic.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15052.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Example 10; Page 84-86; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1043 BP; 212 A; 370 C; 291 G; 170 T; 0 other;

Query Match 28.4%; Score 288.4; DB 24; Length 1043;
Best Local Similarity 58.7%; Pred. No. 1.4e-74;
Matches 606; Conservative 0; Mismatches 396; Indels 30; Gaps 5;

```
Qy        298 ACCAAGACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCC 357
              |||     |  | |  ||||| || |||     | |   || |  |  |  |      |   |
Db        301 ACCGCTCCCGCCAAGAAGACCACGACCGTCGCCAAGGCTTCGACTCCGTCCAACTCGAGC 360

Qy        358 ACCTCTTCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACA 417
              | ||| ||||| |      |||| |   ||   || ||        ||||  ||  . |||
Db        361 AGCTCGTCTTCGGGAAAGTACAGCGCTGTCAGCGGTGGCGCTAGCGGCAACGGCGTCACT 420

Qy        418 ACTCGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACT 477
              || || || ||||| || || || ||||| |||   ||||| || || |||   ||||
Db        421 ACCCGCTACTGGGACTGCTGCAAGGCTTCGTGCTCGTGGCCCGGCAAGGCTAACGTCAGC 480

Qy        478 GGTCCTGTTGACACCTGTGCCTCCAATGG---TATCTCTTTATTAGATGCCAATGCTCAA 534
                |||||  |  ||||   |    | ||    | | || |    ||  |||| || ||
Db        481 TCGCCTGTCAAGTCCTGCAACAAGGACGGCGTCACCGCTCTTAGCGACTCCAACGCCCAG 540

Qy        535 AGTGGTTGTAACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAAT 594
                || || ||||| || ||    | |||||| ||| ||||||| || ||||||||||
Db        541 TCCGGCTGCAACGGCGGCAACTCCTACATGTGCAACGACAACCAGCCATGGGCTGTCAAC 600

Qy        595 GATGAGCTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGG 654
              ||  | || |||||||||||||||||| ||  | |||   |||      |||      | |||
Db        601 GACAACCTTGCTTACGGTTTCGCTGCCGCTGCCATTAGCGGCGGTGGCGAGAGCCGCTGG 660

Qy        655 TGTTGTGGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTT 714
              || ||   ||| |  ||  | |||||||| ||    |   | | |||| ||||||||||
Db        661 TGCTGCTCCTGCTTCGAGCTCACCTTCACCTCCACCAGCGTTGCTGGCAAGAAGATGGTC 720

Qy        715 GTTCAAGTTACCAACACCGGTGGCGATTTAGG---------CTCTAACCACTTTGATTTG 765
              || || || |||||||| || || ||  | ||        |  |  |||||| ||| |
Db        721 GTCCAGGTCACCAACACTGGCGGTGACCTTGGCAGCTCGACCGGTGCCCACTTCGATCTC 780

Qy        766 CAAATGCCCGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCC 825
              || |||||||| || ||||| || |||||||| || ||   |  | || ||||||||||||
Db        781 CAGATGCCCGGCGGCGGCGTCGGCATCTTCAACGGATGCTCGTCCCAGTGGGGCGCTCCC 840

Qy        826 AATGATGGCTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTT 885
              || || ||||||||  |  | || || ||  ||||||| | |   |||||  | || ||
Db        841 AACGACGGCTGGGGCTCGCGCTACGGCGGCATCAGCTCCGCCAGCGACTGCTCGTCCCTC 900

Qy        886 CCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAAC 945
              |||   || || || || || || || ||| | |||||||||||||||||| | || |||
Db        901 CCCAGCGCCCTCCAGGCCGGCTGCAAGTGGCGCTTCAACTGGTTCAAGAACGCCGACAAC 960

Qy        946 CCTACCATGACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGC 1005
              ||  |||||||| |||||| || ||||| ||    ||  | ||  ||    | || |||
Db        961 CCGTCCATGACCTACAAGGAGGTCACCTGCCCCAAGGAGATCACCGCTAAGACCGGATGC 1020

Qy       1006 GAAAGAAAGTAA 1017
                | ||||||
Db       1021 TCGCGCAAGTAA 1032
```

```
RESULT 15
AAV16105
ID   AAV16105 standard; cDNA; 984 BP.
XX
AC   AAV16105;
XX
DT   21-JUL-1998  (first entry)
XX
DE   Fusarium oxysporum non-surface-active endoglucanase gene.
XX
KW   endoglucanase; non-surface-active; cellulase; detergent; cleaning;
KW   performance; stain removal; soften; feel; colour; ss.
XX
OS   Fusarium oxysporum.
XX
FH   Key             Location/Qualifiers
FT   CDS             97..735
FT                   /*tag=  a
FT                   /product= "endoglucanase"
XX
PN   WO9804663-A1.
XX
PD   05-FEB-1998.
XX
PF   25-JUL-1997;   97WO-US13194.
XX
PR   30-JUL-1996;   96US-0023125.
XX
PA   (PROC ) PROCTER & GAMBLE CO.
XX
PI   Boyer SL;
XX
DR   WPI; 1998-130664/12.
DR   P-PSDB; AAW46619.
XX
PT   Detergent composition containing both surface-active and non-surface
PT   active cellulase - softens and improves feel of cotton fabrics
PT   without causing loss in weight or tensile strength
XX
PS   Disclosure; Pages 54-56; 68pp; English.
XX
CC   The sequence is that encoding the endoglucanase enzyme of a non
CC   surface-active cellulase produced by Fusarium. The enzyme can be
CC   used in a detergent composition with a surface-active
CC   cellulase. This combination improves cleaning performance
CC   (maintains colours and removes stains), and softens and
CC   improves the feel of cotton fabrics without causing losses in
CC   weight or tensile strength.
XX
SQ   Sequence 984 BP; 229 A; 270 C; 227 G; 258 T; 0 other;
```

```
  Query Match             21.8%;  Score 221.4;  DB 19;  Length 984;
  Best Local Similarity   63.9%;  Pred. No. 8.2e-55;
  Matches  371;  Conservative    0;  Mismatches  201;  Indels    9;  Gaps    2;

Qy        400 TCTGGCAGTGGTTCCACAACTCGTTATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCT 459
              ||||| || |||   | | |||||  || ||||||||  || ||  ||||||||   |||
Db        151 TCTGGAAGCGGTCACTCTACTCGATACTGGGATTGCTGCAAGCCTTCTTGCTCTTGGAGC 210

Qy        460 GGAAAAGCTTCTGTCACTGGTCCTGTTGACACCTGTGCCTCCAATGGTATCTCTTTATTA 519
              ||||| ||| ||||||  |  |||| |   || ||||      || |  | | |  | |
Db        211 GGAAAGGCTGCTGTCAACGCCCCTGCTTTAACTTGTGATAAGAACGACAACCCCATTTCC 270

Qy        520 GATGCCAATGCTCAAAGTGGTTGTAA---CGGTGGTAATGGTTTCATGTGTAACAACAAC 576
              |  ||||||||   |  ||||||  |    ||||||  || ||      || | |||| ||
Db        271 AACACCAATGCTGTCAACGGTTGTGAGGGTGGTGGTTCTGCTTATGCTTGCACCAACTAC 330

Qy        577 CAACCTTGGGCTGTCAATGATGAGCTCGCTTACGGTTTCGCTGCTGCCTCTATTGCTGGC 636
              || |||||||||||  ||||||||  || ||||||||||||||||  ||    ||  | ||
Db        331 TCTCCCTGGGCTGTCAACGATGAGCTTGCCTACGGTTTCGCTGCTACCAAGATCTCCGGT 390

Qy        637 TCCAACGAAGCTGGATGGTGTTGTGGCTGTTATGAATTGACCTTCACTTCTGGCGCTGCT 696
              |  ||| ||  | |||||  ||||  || ||||  ||||||||||  |||||  | |
Db        391 GGCTCCGAGGCCAGCTGGTGCTGTGCTTGCTATGCTTTGACCTTCACCACTGGCCCCGTC 450

Qy        697 TCTGGAAAGAAGATGGTTGTTCAAGTTACCAACACCGGTGGCGATTTAGGCTCTAACCAC 756
              || |||||||||  | || ||    |||||||| || || ||| | |||   ||||||
Db        451 AAGGGCAAGAAGATGATCGTCCAGTCCACCAACACTGGAGGTGATCTCGGCGACAACCAC 510

Qy        757 TTTGATTTGCAAATGCCCGGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGG 816
              || ||| |    |||||||| ||||| || |||||||||  | |||||   |  || | |
Db        511 TTCGATCTCATGATGCCCGGCGGTGGTGTCGGTATCTTCGACGGCTGCACCTCTGAGTTC 570

Qy        817 GGCGCTCCCAATGATGGCTGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGT 876
              ||      ||| | |  | | || ||     || || ||| ||  |||        || |||
Db        571 GG------CAAGGCTCTCGGCGGTGCCCAGTACGGCGGTATCTCCTCCCGAAGCGAATGT 624

Qy        877 GCCTCTCTTCCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAAC 936
              |       |||     |||  | |  |||||  | ||| |||||  ||||||||  |||||
Db        625 GATAGCTACCCCGAGCTTCTCAAGGACGGTTGCCACTGGCGATTCGACTGGTTCGAGAAC 684

Qy        937 TCTGATAACCCTACCATGACCTTCAAGGAAGTTACCTGTCC 977
```

```
                 | || ||||||  | | |||||  || | |||   || ||
Db        685 GCCGACAACCCTGACTTCACCTTTGAGCAGGTTCAGTGCCC 725
```

Search completed: June 17, 2003, 11:49:11
Job time : 225.839 secs